Custom-Enabled System Architectures for High End Computing

Thomas Sterling
California Institute of Technology
tron@cacr.caltech.edu

Peter Kogge
University of Notre Dame
kogge@cse.nd.edu

Abstract

The US Federal Government has convened a major committee to determine future directions for government sponsored high end computing system acquisitions and enabling research. The High End Computing Revitalization Task Force was inaugurated in 2003 involving all Federal agencies for which high end computing is critical to meeting mission goals. As part of the HECRTF agenda, a multi-day community wide workshop was conducted involving experts from academia, industry, and the national laboratories and centers to provide the broadest perspective on important issues related to the HECRTF purview. Among the most critical issues in establishing future directions is the relative merits of commodity based systems such as clusters and MPPs versus custom system architecture strategies. This paper presents a perspective on the importance and value of the custom architecture approach in meeting future US requirements in supercomputing. The contents of this paper reflect the ideas of the participants of the working group chartered to explore custom enabled system architectures for high end computing. As in any such consensus presentation, while this paper captures the key ideas and tradeoffs, it does not exactly match the viewpoint of any single contributor, and there remains much room for constructive disagreement and refinement of the essential conclusions.

1. Introduction

Two strategies for achieving the high-end computer systems of the future are 1) COTS-based and 2) Custom-enabled computer architecture. COTS-based parallel system architectures exploit the development cost and lead time benefits of incorporating components including microprocessors, DRAM, and interface controllers developed for the mainstream computing market in highly replicated system configurations such as but not limited to commodity clusters. Custom-enabled architectures are designed expressly for the purpose of being integrated in scalable parallel structures to deliver substantially higher performance, efficiency, and programmability than COTS-based systems while requiring lower power and space. Both approaches are likely to lead to Petaflops scale performance prior to the end of this decade, but may exhibit very different operational properties as they are deployed and applied to compute and data intensive applications critical to national security and commerce.

This paper considers the opportunities, technical strategies, and challenges to realizing effective computing performance across the trans-Petaflops regime through possible custom-enabled high-end computer architectures. It reflects the findings of one working group of the HECRTF Workshop (High End Computing Revitalization Task Force)[1,2] conducted in Washington DC from June 16 to 18, 2003 and sponsored by the HEC National Coordination Office. The paper presents the opportunities to be gained, the challenges to be addressed, and a roadmap of progress that may be undertaken to regain US strategic dominance in high performance computing through future advances in custom computer architecture.

2. Custom Enabled Architecture

A custom high end computer architecture is one that has many of the following characteristics:

- its major components are designed explicitly to be incorporated in highly scalable system structures, and operate cooperatively on shared parallel computation to deliver high capability, short time to solution, and ease of programming.
- balanced with respect to rate of computing, memory capacity, and network communication bandwidth.
- exploit performance opportunities afforded by device technologies through innovative structures that are not taken advantage of by conventional microprocessors and memory devices
- incorporate special hardware mechanisms that address sources of performance degradation

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

typical of conventional architectures including latency, contention, overhead, and starvation.
- support improved parallel execution models and assume more responsibilities of global management of concurrent tasks and parallel resources, thus significantly simplifying programmability and enhancing user productivity.

Even though specialized devices are key to the success of the strategy of custom architectures, COTS components are and should be employed where useful when performance is not unduly sacrificed.

2.1 Objectives of HEC Custom Architecture

There are a wide range of possible custom parallel architectures, varying both in strategy and generality. But in all cases, the objectives of their development are to

1) enable the solution of problems we cannot solve now, or of much larger versions of problems that we are currently solving on conventional COTS based systems through dramatic capability improvement,
2) deliver orders of magnitude better performance to cost, size, and power than contemporary COTS systems at the performance scale for which they were designed, and
3) achieve significant reduction to time to solution both through execution performance and enhanced programmability.

2.2 Strategic Benefits

Custom architectures, by their very nature promote a diversity of architecture by relaxing the constraints of system design imposed by conventional COTS microprocessors, and thus open up opportunities for either alternative or point-design solutions to HEC problems that are far more efficient than possible today. Their peak operation throughput and internal communication bandwidth for a given scale system may exceed equivalent attributes of conventional systems by one to two orders of magnitude, overcoming what is often quoted as the road-blocks for current technology. Overall system efficiency may be increased by up to an order of magnitude or more for some challenging classes of applications by means of hardware mechanisms devised expressly for efficient control of parallel actions and resources.

Enhanced programmability is a product of reduced barriers to performance tuning and elimination of many sources of errors, thus simplifying debugging. By efficiently exploiting program parallelism at all levels through superior execution models, efficient control, and sufficient global communication bandwidth, custom architectures exhibit high scalability to solve problems of national importance that may be unapproachable by more conventional means. Custom architectures permit much greater density of computing capability than conventional architectures yielding potentially dramatic reductions in power, size, and cost. Finally, custom architectures may be the only way to achieve sufficient reliability through fault tolerant techniques for systems beyond a certain scale, which may be crucial to realizing systems in the mid to high levels of Petaflops scale.

2. 3 Challenges

In spite of the promise of custom enabled HEC architectures, there are significant challenges to realizing their potential. Foremost among these is that while conventional systems may exploit the economy of scale yielded by the COTS components' mass market, custom architectures, at least initially, will have only a limited market and therefore have fewer number of devices across which to mitigate the development NRE costs. Therefore the benefits achieved through custom design must be able to outweigh the higher per chip price. Also of importance is the longer development lead times that are possible because a larger part of the system needs to be designed from scratch than the COTS based counterparts. Among other consequences of this is that technology refresh is less frequent for custom system architectures.

There exists the challenge of user acceptance resulting from incompatibilities with standard platforms and the need to develop new software environments because of this. Difficulty in porting legacy applications combined with the need for programmer training in the use of the new execution models and tools supported by the custom systems can present further barriers to both users and potential vendors. Finally, initially any new system is unproven in the field and involves real risk to the earliest users. The introduction of any new and innovative custom system must overcome these challenges to be successful.

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

3. Major Conclusions

From the wealth of facts and considerations derived from this important community forum, several key observations emerged from the consensus that should be considered for incorporation in any planning of future HEC procurement and development Federal programs. The most significant among these is presented in the following subsections.

3.1 Advantages

Custom enabled architectures offer significant advantages in performance and programmability compared to COTS based systems of the same scale and deployment time for important classes of applications. Performance advantage of between 10X and 100X is expected through a combination of high-density functionality and dramatic efficiency improvements. Programmability advantage of 2X to 4X is possible both through the elimination or reduction of programmer responsibility for explicit resource management and performance tuning and through advanced execution and programming models providing a reduction of sources of parallel programming errors. Significant advantages in performance to cost are expected to be yielded from the high-density packaging, low power structures, and greater up time from intrinsic fault tolerance mechanisms.

3.2 Near and Medium Term Opportunities

Multiple technical approaches for custom enabled architectures of significant promise have been identified that can be realized in the near and medium term, that with necessary funding will accelerate computing capability, and would permit the US to leap frog foreign competition, thus regaining preeminence in the field. Proof of concept of more than one such innovative architecture is feasible within the next five years, and Petaflops-scale computer systems can be deployed substantially before the end of this decade.

3.3 Strategic Partnerships

Achieving the above opportunities, so important to US national security and commerce, will demand new partnerships among industry, universities, government laboratories, and mission agencies. In order to succeed, such alliances must be coordinated such that the strengths of each institution complement the limitations of the others. Industry provides the principle skills and resources to manufacture complex computing systems, but lacks the motivation to explore high-risk concepts. University research groups devise and investigate innovative directions that could lead to future system types, but lack the resources or organization to carry them through to useful form. The National laboratories have the expertise of using the largest high-end computers for major applications of importance to the national welfare, but do not develop the computing engines that they use. And the Federal agencies have both the requirements and the resources to enable future useful systems to be invented, evaluated, and if warranted deployed, but have at best only limited abilities to help steer commercial technologies to niche markets such as HEC. No one side of the community can realize the opportunities of future custom architecture alone and a new class of peer-to-peer partnering relationship is necessary to restart the HEC research pipeline with new ideas, faculty, and graduate students.

3.4 Funding Culture

The current funding culture is incapable of enabling, let alone catalyzing, the revitalization of the HEC industry and research community. The narrow short-term specifications, limited (even single year) time frames, woefully inadequate budget levels, insufficient guarantees to industry as friendly customers, and conflicting objectives across agencies has dissipated the means and will of the HEC community to attempt to provide anything but incremental advances to conventional COTS based systems, leaving future innovation to foreign suppliers. The resulting soft money mentality has largely eliminated research incentive and been disruptive to national initiative compared to the Japanese programs, which produced the Earth Simulator and is likely to deliver the first Petaflops scale computer within the next two years.

3.5 Innovation in System Software and Programming Environments

While it is the finding of this expert panel that the exploitation of custom architectures devised for the explicit purpose of scalable parallel computing is imperative to achieve the full potential of the foundation technologies, it is also clear that this alone is insufficient in meeting the goal. Both system software and programming environments must be developed that support and exploit the capabilities of the custom architectures. System software must be developed to provide the dynamic resource management anticipated by many of these

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

architectures to improve performance efficiency and remove much of the burden from the programmers. Programming environments must be developed that capture and expose intrinsic algorithm parallelism for greater performance, and provide high level constructs to eliminate low level and error prone detail to minimize application development time. In addition, effective software means must be provided to enable rapid porting of legacy applications and libraries to maintain continuity of the user software base. The creativity of future software and programming models must match the creativity in custom HEC architecture. The required investment in software development is likely to exceed that of the custom architecture by at least a factor of 4X (some would estimate it at 10X).

3.6 Applications Requirements Characterization

Future Petaflops scale architectures, whether custom or COTS based, will run applications of substantially larger scale and complexity than those performed on current generation MPPs and clusters. In some cases, entirely new applications and/or algorithms not even attempted in the current environment may become important users of future systems. Therefore, there is little (almost none) quantitative characterization of the actual system requirements of these future systems. Against the expected sources of user demand for such systems, it has not been determined with any certainty what the resource needs will be for memory capacity, network bandwidth, task parallelism control and synchronization, I/O bandwidth, and secondary storage capacity.

3.7 Basic Research for End of Moore's Law

It is expected that by 2020 the exponential growth in silicon semiconductor device density, usually attributed to Moore's Law, will have terminated ("flatlined") due to a number of causes, and that significant reduction in the rate of performance growth due largely to silison technology may be experienced as early as 2010 or shortly thereafter. Beyond that time frame, continued growth in system performance will be derived primarily through brute force scale, advances in custom computer architecture, and incorporation of exotic technologies. In this last case, architecture advances will be required to best assimilate such novel materials and adapt computing structures to their behavioral properties. Therefore, it is necessary that basic research be initiated in the near future for custom architectures that will be prepared for the end of Moore's Law and the introduction of alien technologies and models. It is expected that there is the potential for significant trickle-back to silicon based semiconductor system architecture, even prior to the time when such innovations in architecture will become imperative.

4. Technical Directions for Future Custom Architectures

In spite of a period of limited funding for HEC computer architecture research, a number of paths have emerged that hold the real potential for one to two orders of magnitude advantage in several critical dimensions with respect to conventional architecture and practices using current or near term technologies. Further, it is clear that these gains will continue to prevail through architectural and complementing system software means through at least the end of the decade, benefiting proportionally from enhanced semiconductor technology improvements governed by Moore's Law. This section documents key technical opportunities and potential advances that will be delivered by custom architecture research should such work be adequately funded.

4.1 Fundamental Opportunities Enabled by Custom Architecture

Custom architecture uniquely is able to exploit intrinsic and fundamental opportunities implicit in available or near term underlying technologies through innovative structures and logical relationships. Some of the most important are suggested here:

4.1.1 Function Intensive Structures

The low spatial and power cost of VLSI floating point arithmetic and other functional units permits new structures incorporating many more such elements throughout the program execution and memory service components of future parallel system architectures. Organizations comprising 10X to 100X more functional units within a corresponding scaled HEC system is feasible in the near term, assuming logical control and execution models are devised that can effectively coordinate their operation.

4.1.2 Enhanced Locality – Increasing Computation to Communication Demand

Communication is a major source of performance degradation, whether global across a system or local across a single chip. It is also a major

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

source of power consumption. Custom architectures present the opportunity through innovative structures to address both scales of communication, even to a significant degree in some cases, to significantly increase the computation to communication ratio.

4.1.3 Exceptional Global Bandwidth

Custom HEC system architectures are distinguished from their COTS-based counterparts by interconnecting all elements of the distributed system with exceptional global bandwidth and at relatively low latency. In so doing, custom architectures can significantly reduce several sources of performance degradation typical of conventional systems, including: contention for shared communication resources, delay due to transit time of required remote data, and overhead for managing the global network. Depending on the system used as a basis for comparison, improvements can easily exceed 10X and approach 100X. Such global bandwidth gains not only improve performance, it can greatly enhance the generality of high end systems in supporting a wide range of application/algorithm classes, including those which are tightly coupled, are communication intensive, and involve substantial synchronization. Increased bandwidth also improves architecture scalability.

4.1.4 Architectures that Exploit Global Bandwidth

Bandwidth alone, although a dominant bounding condition on system capability is insufficient to guarantee optimal global performance. Custom architectures must in addition incorporate means to support many outstanding in-flight communication requests simultaneously, and if possible permit out of order delivery. This requires a combination of methods including special lightweight mechanisms for efficient management of communication events and higher-level schema for representing and managing a high degree of computation parallelism. With high concurrency of demand and low overhead of operation, the raw exceptional capacity of custom global interconnection technology and network structures may be effectively exploited.

4.1.5 Efficient Mechanisms for Parallel Resource Management

A repeated requirement governing many aspects of HEC system operation is efficient mechanisms for the management of parallel resources and the coordination of concurrent tasks, especially at the fine grain level. Fine grain parallelism, which is crucial to scalability of future Petaflops systems, can only be exploited if the mechanisms responsible for their operation and coordination are fast enough such that the temporal overhead does not overwhelm the actual useful work being performed. Custom HEC architecture has the unique advantage of being able to incorporate such hardware supported and software invoked mechanisms employed for global parallel computation.

4.1.6 Advanced ISA

To facilitate the control of widely distributed and highly parallel HEC system architectures, the semantics of parallel operation needs to be reflected by the instruction set microarchitecture. This is only possible through custom system and microarchitecture design. Otherwise, all responsibilities of managing concurrency of resources and tasks must be emulated through software, often requiring egregious use of synchronization variables and the overhead that entails. There are also classes of operations, which while not particularly important to general commercial computation, and therefore not usually found as part of COTS microprocessor instruction sets, nonetheless can be very important to scientific/technical computing as well as to the mission critical computations of defense related agencies. Custom architecture may provide optimized instructions for these and other purposes that will never be available from COTS-based systems.

4.1.7 Execution Models that Facilitate Compiler/Programmer Application

Beyond the specifics of instructions and components, the overall operational properties of a highly scalable, efficient, and programmable parallel computing system is governed by an abstract schema for defining the relationships among the actions to be performed and data upon which they are to operate. In an actual parallel computer, such a representation formalism is manifest as an execution model that determines the emergent behavior of the system components in synergy with support of the user application. The execution model establishes the principles of control and is supported by the instructions, the mechanisms, and the system structure. It enables the compiler and programmer to effectively employ the capabilities of the resources comprising the system. A COTS based system is extremely limited in the choices of execution models because they fail to provide the needed underlying functionality.

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

4.2 Examples of Innovative Custom Architectures

A number of concepts of innovative custom architectures were identified by the panel and their specific characteristics and advantages examined. Each of these incorporates structures and strategies that exploit one or more of the potential opportunities previously discussed. An incomplete set of examples of possible innovative custom architectures is presented in this section.

4.2.1 Spatially Direct Mapped Architecture

An important strategy to achieve high density of functional units, low latency between successive operations, high computation to communication, and low power consumption is to enable structures of functional units and their interconnection paths to closely match the intrinsic control flow and data flow of the application kernel computation. There are several ways to do this, and the different strategies vary in their flexibility and efficiency. The "spatially direct-mapped architecture", also referred to as "adaptive logic" or "reconfigurable logic" comprises an array of logic, storage, and internal communication components the interconnection of which may be programmed and changed rapidly, sometimes within milliseconds. The goal (and reason for the term "spatial") is to allow us to compile not to a temporal sequence of ordered instructions, but to a spatial surface through which data flows.

4.2.2 Vectors

Vector processing exploits pipelining of logic functions, communication, and memory bank access to exploit fine grain parallelism for efficient high performance computation. It provides a class of efficient fine grain synchronization, the potential of overlap of communication with computation, and reduced instruction pressure. While best at exploiting dense unit stride accesses, additional mechanisms permit rapid gather scatters across more widely varying access patterns. The vector model has been successfully exploited since the 1970s but new implementation strategies are emerging that will extend its capability through innovative architectures.

4.2.3 Streaming Architecture

Streaming architecture is being proposed as an innovative strategy for providing a very high-density logic architecture with full programmability. Wide and deep arrays of arithmetic functional units are interconnected with intermediate result data transiting through the array driven by a software/compiler controlled communication schedule. Very high computation to communication can be achieved for certain classes of algorithms, exhibiting high computation rate and low power.

4.2.3 Processor in Memory Architecture

Processor in Memory (PIM) architecture also exploits a high degree of logic density but in a form and class of structure very different from those of vectors, streaming, and spatially direct mapped architectures. Instead, PIM merges arithmetic logic units with memory such that the logic is tightly coupled with the memory row buffers. With access to the entire row buffer, wide ALUs can be employed to perform multiple operations on different data within a single memory block at the same time. The total memory capacity of a memory chip may be partitioned into many separate units potentially exposing > 100X memory bandwidth at low latency for data intensive low/no temporal locality operation.

4.2.4 Special Purpose Devices

Special purpose devices (SPD) are hardwired computational structures that are optimized for a particular application kernel. They take advantage of the same mapping attributes as spatially direct mapped (reconfigurable) architectures. But they are able to exploit very high-speed technology and provide much greater logic density to deliver significantly greater performance per unit area and lower power per computing action. SPDs such as systolic arrays have a long history of development and are particularly useful for post sensor and streaming data applications. The world's fastest (unclassified) computer, Grape-6, is of this type and it is likely that the first Petaflops scale computer will be a derivative of this architecture, to be deployed within the next two years. An important limitation of SPDs is, as their name implies, that they are limited in the range of computations that any one of them can perform.

4.3 Enabling and Exploiting Global Bandwidth

Custom architectures may be distinguished from their COTS-based counterparts in part by enabling exceptional global bandwidth and its effective exploitation. Global networks for future HEC custom systems exhibiting order of magnitude greater bi-section bandwidth than conventional systems, and may employ advanced technologies including high speed signaling for both optical and electrical

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

channels as well as heterogeneous mixes, possibly using VCSEL arrays. Optical switching and routing technologies will also be employed but it should be noted that routing and flow control are already nearing optimal capability. High bandwidth, high-density memory devices might also facilitate fast communications (the panel notes that the current generation of commodity memory devices do not provide anywhere near the external bandwidth that they could be capable of using existing advanced signaling protocols, or that they already have available within the chip from the memory arrays.).

From these base technologies, advanced network structures may be created. High radix networks organized in non-blocking bufferless topologies will be deployable within a few years using a combination of hardware congestion control and compiler scheduled routing strategies. A number of processor architecture advances are key to providing a sufficient traffic stream to utilize these future generation enhanced networks for high efficiency. Within the processor control of fine-grain parallelism, architectures incorporating streams, vectors, and multithreading provide the large numbers of simultaneous in-flight access requests per processor to make good use of such enhanced global network resources. Global shared memory and low overhead message passage mechanisms make lightweight packets feasible, providing additional concurrent global network traffic. Other techniques such as prefetch and prestaging mechanisms as well as other methods of augmenting microprocessors to enhance additional requests also contribute to the parallelism of communication and the effective exploitation of global bandwidth.

4.4 Enabling and Exploiting Function Intensive Structures

Among the foremost opportunities for custom architecture are two related ones: the tremendous potential expansion of arithmetic functional units on a per die basis to increase peak floating point bandwidth by one to two orders of magnitude, and more importantly greatly enhancing processor internal bandwidth and control locality. Spatial computation via reconfigurable logic is one such architectural method. Streams that capture physical locality by observing temporal locality is another. New methods embodied at the microarchitecture level promise to enhance both locality and scalability of vectors. Processor in memory captures spatial locality via high bandwidth local memory with low latency and exploits high logic capability by enabling many active data/logic paths on the same chip. Chip stacking may further increase local bandwidth and logic density. General techniques of software management of deep and explicit register and memory hierarchies may lead to further exploitation of high logic density.

4.5 Efficiency through Custom Mechanisms

Efficiency of execution and scalability demand the ability to exploit fine grain tasks and lightweight communication. Custom architectures provide a unique opportunity through the design of hardware mechanisms to be incorporated in the processor, memory, and communications elements. Such mechanisms can provide high-speed means of synchronization, context switching, global address translation, message generation, routing, switching, acquisition, and interpretation. Fast methods of memory management, cache handling, security in a global parallel system can greatly reduce factors contributing to lower efficiency.

4.6 Execution Models

In order to effectively exploit the capabilities of custom architectures as described in the previous sections, execution models must be devised that govern the control of the global parallel system in response to the computational demands of the user applications. A good model should expose parallelism to the compiler and system software, provide explicit performance cost models for key operations, not constrain the ability to achieve high performance, and provide an abstract logical interface for ease of programming. While no single execution model for, and supported by, custom architectures was selected, potential elements of such a future model of computation were identified based on the classes of parallel architectures being considered. The spatial direct mapped hardware approach suggests its own paradigm, although in the limit it could efficiently emulate many different such models. Low overhead synchronization mechanisms open up the prospect for a rich array of parallel constructs and the potential of new memory semantics. With the prospect of PIM enabled architectures, these can be further advanced along with additional fundamental constructs such as message-driven computation, traveling threads, active pages, and so on. Streams and threads extend the space of lightweight efficient parallelism that both support and are supported by future execution models. Such models must distinguish between local (uniform access) and global (non-uniform access) memory structures and access policies. A first good step is the programming models represented by Co-Array Fortran and UPC. But far

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

more sophisticated execution models will be required in order to fully exploit the potential of promising new custom architectures.

5. Discussion of Open Issues

5.1 The Relationship between Programming and HEC Architecture

High-End Computing is general purpose. The applications demanding HEC performance will use a wide variety of algorithms and data structures, both known and yet to be developed. Moreover, large simulations will frequently involve different sorts of models for different components, and different approaches for several time scales important to a calculation. Many applications will not match the massively parallel, data parallel, or MPI models that are most efficiently supported by today's HEC machines. Therefore, future HEC architectures will need to support Programming-in-the-Large. It must be possible to put large programs together by combining components separately implemented. Component interfaces should be simple and independent of the mechanisms used internally. Programs for large simulations will often be so large that compiling in one step is not practical. Some of the proposed architecture ideas are not general purpose and/or do not explain how Programming-in the Large would be supported. A large-scale HEC must support simultaneous execution of multiple jobs for different users with security. This is necessary to make possible efficient utilization of the expensive equipment. Any complete architecture proposal must indicate how this will be supported. Application people have requested both programmer/compiler direction of resource allocation, and dynamic resource management at run time. For dynamic resource management, hardware support is essential, and must be included in any complete proposal. Also, a global shared address space is essential for a sufficiently general dynamic resource management implementation.

5.2 The Role of Universities

In the quest to dramatically regain US leadership in high end computing, universities provide a critical resource. Academic institutions are a major source of innovative concepts and long-term vision. They are the major source for keeping the research pipeline full, in part as they provide the students who are engaged in formulating and testing new ideas, and developing the skills required to carry them out. Universities have demonstrated their proficiency at conducting early simulations of conceptual hardware and software systems, and are a major development facility for prototype tools. While it is difficult to produce leading-edge integrated circuits, universities are just about the only venue of implementing 1st generation prototypes of novel concepts. It is noted that a general trend in computer science education today is that students are no longer commonly exposed to massive parallelism in particular, and that there is a significant decline in students of parallel computer architecture in general, as well as an atrophying interest in high end computing. Universities are only part of the solution and are limited in some ways. They do not do well in taking the work beyond the early research stage (there have been notable exceptions; e.g. Unix-BSD) to the realm of robust products. Due to the ephemeral tenure of student engagement, teams are a challenge to retain; this aggravated by the difficulties imposed by soft money and the uncertainties of funding. This last issue should be addressed as part of the overall strategy devised by the Task Force.

6. Road Map

A general time line of possible advances for innovative custom architecture research and development can be projected based on assumptions of sufficient and sustained funding using the concepts presented earlier as a basis for technical exploration. To this end, three epochs of five years each are considered beginning in FY05, the first fiscal year for which funding derived from this initiative may be anticipated. It is recognized that only funding for the first five-year period will be determined initially. But planning, even for this phase, requires a long-term perspective and vision in order to identify early basic research activities required in preparation for future conditions such as the end of Moore's Law or the introduction of new execution models or technologies.

6.1 5 Years: FY05 – FY09

6.1.1 Deployable

Within the first five years, specific custom architecture elements will be deployable in HEC systems delivered to government agencies prior to the end of this decade. Advances in network technology can provide a new generation of high bandwidth interconnection links, drivers, and routers exhibiting 10X or more bandwidth, and latencies below 1 microsecond across very large systems. New high bit rate wire channels, optical fiber interconnects, and

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

high radix routers together can deliver critical global bandwidth gains over conventional means in real world systems by the end of the decade. Symmetric multithreaded architectures will become ubiquitous within the next five years. Spatial direct mapped architectures (i.e. adaptive logic) can be deployed as well for friendly customers with significant advances in software and compilation strategies accomplished in this time frame.

6.1.2 Prototypes and Experiments

Several advanced architecture concepts identified previously in this report could be developed and prototyped within the first five years of a new initiative. Such experiments would permit evaluation at sufficient depth to determine which specific concepts warrant investment during the second phase to bring them to a maturity level sufficient for deployment.

6.1.3 Basic Research and Exploratory Studies

Beyond the continued current trends of silicon based semiconductor technology dictated by Moore's Law, innovations in device structures and technologies, and radical changes in architecture and execution models will require fundamental basic research to be initiated within the first five years of the new program. The early exploratory studies will develop inchoate concepts and push the edge of the envelop to provide the mission agencies with alternatives in order to avoid plateauing of capabilities in the early part of the second decade. Such research could include new computational models, nanotechnology, quantum dots, cellular automata, amorphous computing, and continuum computer architecture. New resilient fault tolerant and decentralized control (e.g. distributed agents) architectures would yield robust systems at scales where single point failure modes would limit sustained up times to seconds. New approaches to compilers and runtime systems as well as scalable I/O and operating systems would be undertaken, and would be strongly influenced by work on novel programming models.

6.2 10 Years: FY10 – FY14

The second five-year epoch of a new funded initiative in custom computer architecture could prove to be an explosive renaissance in system design, capability, robustness, and usability. All prior prototypes and experiments that had demonstrated viability during the first five-year phase of the program could be deployed at mission agency sites. Such systems would provide sustainable performance for general applications in the 10 to 100 Petaflops performance regime and exhibit competitive recurring cost compared to conventional techniques, while delivering far superior operational attributes for at least many important agency applications, assuming they were properly funded. Virtually all of prior technology opportunities that were developed in the first phase will be deployable by the second phase in real world HEC systems. But initial adoption of such systems will be limited by drastic changes required in execution and programming models, although methods for transferring legacy codes to such radical systems will be a continued area of important research. Infrastructure between academia and industry will have to be established to encourage and enable the transfer of research results and their incorporation in deployed hardware and software systems.

6.3 15 Years: FY15 – FY19

With silicon scaling at sunset, systems developed and deployed during the latter part of the 2nd decade will exploit revolutionary techniques in circuits, packaging, architecture, and software strategies. These truly revolutionary custom architectures will mesh with the end of the silicon roadmap and new non-silicon technologies that will have been proven during the previous phases of the program. As these exotic systems are prototyped and deployed, alien in form and function but capable of near Exaflops performance, entirely new software environments for resource management and programming will have been devised and will be developed during this period. Ironically, these systems delivering a hundred thousand times the performance of today's HEC systems may be smaller, consume less power, and take up much less space than even today's Teraflops systems.

7. Summary and Conclusions

Custom-enabled HEC architecture provides a vital alternative to conventional COTS-based system design by enabling the exploitation of potential advantages intrinsic to available and near term technologies, but demanding innovative hardware structures and software management models and methods. In many important metrics, custom architectures may deliver between 10X and 100X advantage over conventional COTS based systems employing equivalent semiconductor technology including peak and sustained performance, performance to cost, power, size, and reliability.

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE

Custom architectures may efficiently support advanced execution and programming models that will both deliver superior sustained performance and greatly facilitate programmability, thus enabling systems of exceptional productivity for federal agency mission driven applications. It is imperative that research in advanced custom scalable HEC architecture be sponsored at an accelerated and continuous level to regain US leadership in the field of HEC architecture and provide the tools to secure dominance in this strategically critical technology for national security and commerce. To a large extent, the students we train in the first epoch will be the ones doing this final epoch work, and it is crucial that we give them the mind set, tools, and funding to be able to set a truly innovative and aggressive research plan 15 years from now.

References

[1] Workshop on The Roadmap for the Revitalization of High-End Computing. Daniel A. Reed (Ed.), Computing Research Association, Washington, DC, 2003.

[2] High-End Computing Revitalization Task Force (HECRTF). Federal Plan for High-End Computing. Report to the Executive Office of the President, Office of Science and Technology, Washington, DC, 2004.

Proceedings of the Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA'04)
1527-1366/04 $ 20.00 IEEE